om

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/020,895	12/19/2001	Soren Riis	042933/301043	8838

826 7590 02/21/2007

ALSTON & BIRD LLP
BANK OF AMERICA PLAZA
101 SOUTH TRYON STREET, SUITE 4000
CHARLOTTE, NC 28280-4000

EXAMINER	
NG, EUNICE	

ART UNIT	PAPER NUMBER
2626	

SHORTENED STATUTORY PERIOD OF RESPONSE	MAIL DATE	DELIVERY MODE
3 MONTHS	02/21/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

Office Action Summary	Application No.	Applicant(s)
	10/020,895	RIIS ET AL.
	Examiner	Art Unit
	Eunice Ng	2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *14 September 2006*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *13-16,18-25 and 27-40* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *13-16,18-25 and 27-40* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

Continued Examination Under 37 CFR 1.114

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 11/30/06 has been entered.

Response to Amendment

2. In response to the Office Action mailed 03/31/06, Applicants have submitted an Amendment, filed 11/30/06, canceling claim 26, amending claims 13, 15, 22-25, 27, 29 and 31, adding new claims 33-40, without adding new matter, and arguing to traverse claim rejections.

3. Claim 24 has been labeled as “Original” but is --Currently Amended--.

Response to Arguments

4. Applicant's arguments with respect to claims 13, 15, 22-25, 27, 29 and 31 have been considered but are moot in view of the new ground(s) of rejection, below.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 13, 14, 18, 21, 25, 27-29 and 33-40 are rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al. ("Measure of local speaking-rate for automatic speech recognition," published May 13, 1999) in view of Gandhi et al., US Patent No. 5,687,287.

Regarding claim 13, Russell et al. teach a speech recognition system in which an utterance to be recognized is represented as a sequence of phonetic segment models (see abstract, discussing "phone-level" speaking and estimation) in which a transition probability represents the probability of the occurrence of a transition between the models (see lines 4-5, "N-state HMM... transition probability" under "ROS compensation"), comprising means (a speech recognizer) for:

estimating the number of phonetic segments in the utterance to be recognized (see lines 1-2 under "Phone-level measures of ROS" describing a measure of "phones-per-second" (or phonetic segments) in a sentence (which necessarily includes an utterance); Russell estimates the number of phones-per-second, which inherently teaches estimating the number of phonetic segments in the utterance); and

biasing the transition probabilities in dependence on the length of the utterance (see lines 9-10 under "ROS compensation," which discuss the state transition probabilities "scaled for fast speech," implying dependence on length).

Russell et al. does not explicitly teach estimating the number of phonetic segments in the word to be recognized; and biasing the transition probabilities in dependence on the estimated number of phonetic segments in the word. However, this feature is well known in the art as

evidenced by Gandhi et al. in col. 7, ll. 25-33, which teaches segmenting test utterances into words and computing a duration [number of phonetic segments] normalized likelihood score for each word in the input string.

It would have been obvious for one of ordinary skill in the art at the time the invention was made to bias the transition probabilities in dependence on the estimated of phonetic segments in a word instead of an utterance because so that recognition performance for relatively long and/or short words can be improved.

Regarding claim 29, Russell et al. teach a method of speech recognition in which an utterance to be recognized…transition between models, the method comprising biasing the transition probabilities in dependence of the number of phonetic segments in the utterance (see lines 1-2 under "Phone-level measures of ROS" describing a measure of "phones-per-second" (or phonetic segments) in a sentence (synonymous with an utterance); while Russell estimates the number of phones-per-second, it inherently teaches estimating the number of phonetic segments in the utterance; see also p. 1, col. 2, line 11 of "Experimental results", which teaches "K, for each occurrence of a phone symbol in the test set [number of phonetic segments in the utterance]"; Fig. 3 shows the correlation p_k for PRROS estimation window sizes K between 1 and 20, "the identities and endpoints of $p_1, \ldots, p_k$…can be estimated during recognition through partial traceback, and used to adapt the self-transition probabilities according to eqn. 2 throughout an utterance (p. 2, col. 1)," suggesting biasing the transition probabilities in dependence on the number of phonetic segments in the utterance).

Russell et al. does not explicitly teach estimating the number of phonetic segments in the word to be recognized; and biasing the transition probabilities in dependence on the estimated number of phonetic segments in the word. However, this feature is well known in the art as evidenced by Gandhi et al. in col. 7, ll. 25-33, which teaches segmenting test utterances into words and computing a duration [number of phonetic segments] normalized likelihood score for each word in the input string.

It would have been obvious for one of ordinary skill in the art at the time the invention was made to bias the transition probabilities in dependence on the duration a word instead of an utterance so that recognition performance for relatively long and/or short words can be improved.

Regarding claim 14, Russell et al. teach wherein the biasing means comprise means for applying a transition bias to each of the transition probabilities between a plurality of phonetic segment models (see lines 18-21 under "ROS compensation").

Regarding claim 18, Russell et al. teach wherein the estimating means comprises a speaker specific rate of speech estimator (see Abstract).

Regarding claim 21, Russell et al. teach wherein the transition bias is set in response to the result of the estimating means (see lines 6-10 under "ROS compensation," which discuss a rate of speech compensation which scales (or biases) the state transition probabilities according to the speaker specific rate of speech).

Regarding claim 25, Russell et al. teach wherein the, or each, phonetic segment comprises a phoneme (see lines 1-2 under "Phone-level measures of ROS" describing "phone-level" measures wherein a "phone" is a sound unit of speech also known as phoneme, or allophone, which is predictable phonetic variant of a phoneme).

Regarding claim 27, Russell et al. teach wherein an utterance to be recognized is represented as a sequence of phonetic segment models in which a transition probability represents the probability of occurrence of a transition between the models (see lines 1-5, "N-state HMM... transition probability" under "ROS compensation"), comprising:

a phonetic segment estimator arranged to output an estimate of the number of phonetic segments in the utterance (see lines 1-2 under "Phone-level measures of ROS," wherein the utterance is a sentence; while Russell estimates the number of phones-per-second, it inherently teaches estimating the number of phonetic segments in the utterance); and

a processing module for applying a transition bias to the transition probability in response to the output of the estimator (see lines 6-10 under "ROS compensation," which discuss a rate of speech compensation which scales (or biases) the state transition probabilities according to the speaker specific rate of speech).

Moreover, on p. 1, col. 2, line 11 of "Experimental results", Russell teaches "*K*, for each occurrence of a phone symbol in the test set [number of phonetic segments in the utterance]" ; Fig.3 shows the correlation p_k for PRROS estimation window sizes K between 1 and 20, "the identities and endpoints of $p_1, ..., p_k$...can be estimated during recognition through partial

Regarding claims 33, 35 and 37, Russell et al. does not explicitly teach, but Gandhi et al. teaches performing word recognition for the word on an individual basis based on the biased transition probabilities (col. 7, ll. 25-56). It would have been obvious for one of ordinary skill in the art at the time the invention was made to modify the teaching elements of Russell with Gandhi because biasing the transition probabilities in dependence on the estimated of duration of a word instead of an utterance allows the recognition performance for relatively long and/or short words to be improved.

Regarding claim 39, Russell and Gandhi et al. teach comprising: receiving a word to be recognized represented as a sequence of phonetic segment models in which a transition probability represents the probability of the occurrence of a transition between the models; and biasing the transition probabilities in dependence of the number of phonetic segments in the word (as discussed in the rejections of claims 13, 27 and 29, above).

Gandhi et al. also teaches performing word recognition for the word on an individual basis based on the biased transition probabilities (col. 7, ll. 25-56). It would have been obvious for one of ordinary skill in the art at the time the invention was made to modify the teaching elements of Russell with Gandhi because biasing the transition probabilities in dependence on the estimated of duration of a word instead of an utterance allows the recognition performance for relatively long and/or short words to be improved.

traceback, and used to adapt the self-transition probabilities according to eqn. 2 throughout an utterance" (p. 2, col. 1), suggesting biasing the transition probabilities in dependence on the number of phonetic segments in the utterance).

Russell et al. does not explicitly teach estimating the number of phonetic segments in the <u>word</u> to be recognized; and biasing the transition probabilities in dependence on the estimated number of phonetic segments in the <u>word</u>. However, this feature is well known in the art as evidenced by Gandhi et al. in col. 7, ll. 25-33, which teaches segmenting test utterances into words and computing a duration [number of phonetic segments] normalized likelihood score for each word in the input string.

It would have been obvious for one of ordinary skill in the art at the time the invention was made to bias the transition probabilities in dependence on the duration of a word instead of an utterance so that recognition performance for relatively long and/or short words can be improved.

<u>Regarding claim 28</u>, Russell et al. teach a portable communications device including a speech recognition system (see line 16 under "experimental procedure," describing the use of a "DERA ASTREC speech recognizer," which is a state-of-the-art reconfigurable continuous automatic speech engine (or system) from The Defense Evaluation and Research Agency, which is suitable for deployment in command-and control direct voice input applications in a wide range of existing commercial markets (e.g. automotive, telephone-based IVR systems, TV control, etc.) and has already been trialed in a range of applications (e.g. European Fighter Aircraft), which reads on implementation in portable communication devices).

Regarding claims 34, 36, 38 and 40, Russell et al. does not explicitly teach, but Gandhi et al. teaches performing word recognition for each word in a multiword sentence based on a biased transition probability determined separately for each corresponding word in the sentence based on the estimated number of phonetic segments in each corresponding word. However, this feature is well known in the art as evidenced by Gandhi et al. in col. 7, ll. 25-33, which teaches segmenting test utterances into words and computing a duration [number of phonetic segments] normalized likelihood score for each word in the input string.

It would have been obvious for one of ordinary skill in the art at the time the invention was made to bias the transition probabilities in dependence on the estimated of duration of a word instead of an utterance so that recognition performance for relatively long and/or short words can be improved.

7. Claim 19 is rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al. and Gandhi et al. in view of James et al. ("A Fast Lattice-Based Approach to Vocabulary Independent Wordspotting," ICASSP 1994, pp. 377-380).

Russell and Gandhi et al. fail to teach a system wherein the estimating means comprises a Free Order Viterbi decoder. However, Viterbi decoders are well known in the field of speech recognition as evidenced by James et al., which disclose implementing a Free-Order Viterbi decoder (a null-grammar phone network, see page I-379, lines 14-15 of section 3.3).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teaching elements of Russell and Gandhi et al. with those of James et al., because James et al. teach that this would increase flexibility by being able to search for any

word and speed of retrieval (see page I-377, sixth paragraph, lines 1-5; see also US Patent 6,073,095 to Dharanipragada et al. which references this publication in the "Prior Art" section of column 1).

8. Claim 20 is rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al. and Gandhi et al. in view of Bergstrom et al., US Patent No. 5,737,716 (filed Dec. 26, 1995).

Russell and Gandhi et al. fail to teach a system wherein the estimating means comprises a neural network classifier. However, this feature is well known in the art as evidenced by Bergstrom et al., which disclose a neural network controlled speech analysis processor that includes a neural network which manages speech characterization, encoding, decoding, and reconstruction methodologies, reading on a neural network classifier (see abstract).

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify the teaching elements of Russell and Gandhi et al. with those of Bergstrom et al., because Bergstrom et al. teach that this would "provide for rapid development, improved classification accuracy, improved speech analysis and speech synthesis architectures, and improved immunity to interference when trained with appropriate characteristic features" (see column 3, lines 15-19).

9. Claims 15, 16, and 30 are rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al. and Gandhi et al. in view of Gupta et al. (US Patent No. 5,390,278).

Regarding claims 15 and 16, Russell and Gandhi et al. fail to teach a system operable to recognize words from a recognition vocabulary, wherein the transition bias is calculated as the

transition bias which maximizes recognition performance on a validation data set which represents, or has the same vocabulary as, the recognition vocabulary.

However, this procedure would have been obvious to one of ordinary skill in the art at the time the invention was made given the invention by Gupta et al. Gupta et al. teach transition probabilities calculated, with "the one resulting in the best score" stored (see column 17, line 48-49), suggesting choosing a transition bias which maximizes recognition performance, and a validation data set representing, or having the same vocabulary as, the recognition vocabulary (see column 12, lines 45-49 and column 14, lines 21-23).

Russell et al. does not explicitly teach a speech recognition system operable to recognize <u>word</u> from a recognition vocabulary. However, this feature is well known in the art as evidenced by Gandhi et al. in col. 7, ll. 25-33, which teaches segmenting test utterances into words. It would have been obvious for one of ordinary skill in the art at the time the invention was made to recognize words because Gandhi teaches that biasing the transition probabilities in dependence on the duration of a word allows recognition performance for relatively long and/or short words to be improved.

<u>Regarding claim 30</u>, Russell and Gandhi et al. fail to teach comprising decoding the sequence of phonetic segment models after application of the transition bias. However, this procedure would have been obvious to one of ordinary skill in the art at the time the invention was made given the invention by Gupta et al.. Gupta et al. suggest decoding the sequence of phonetic segment models after applying a bias (see Abstract and column 18, first paragraph; decoding is done by the A* search method as illustrated in Fig. 12a., element 418). Motivation

for the combination would be to save the unnecessary decoding before the application of the transition bias, wherein the transition bias improves recognition.

10. Claim 31 is rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al. and Gandhi et al. in view of Gupta et al. (US Patent No. 6,138,095).

Russell and Gandhi et al. fail to teach comprising decoding the sequence of phonetic segment models without the application of transition bias (as specified in the rejection of claim 14, Russell et al. teaches only a transition bias) and normalizing the resulting scores by a contribution proportional to the transition bias.

However, this procedure would have been obvious to one of ordinary skill in the art at the time the invention was made given the invention by Gupta et al.. See column 3, lines 9-24 and column 3, line 66 through column 4, line 2 of Gupta et al. which discloses normalizing rejection thresholds and likelihood ratios (similar to resulting scores) by the magnitude of a null hypothesis probability (similar to transition probabilities). Motivation for the combination would be to simplify processing, in the case where the transition biases are too large, too small, or not integral numbers.

11. Claim 32 is rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al. in view of Gandhi et al. and Gupta et al. (US Patent No. 6,138,095), and further in view of Ueyama et al. (US Patent Application Publication 2001/0056346).

Russell, Gandhi, and Gupta et al. fail to teach comprising calculating the transition bias in parallel with the decoding of the sequence of phonetic segment models. However, this procedure

is well known in the art as evidenced by Ueyama et al., which disclose computing the output probabilities (synonymous to a transition probability) of acoustic models in parallel to decoding of speech parameters (synonymous with a sequence of phonetic segment models). See paragraph [0095]. Motivation for the combination would be to save time.

12. Claims 22-24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Russell et al. in view of Gandhi et al. and Schwartz et al. (US Patent No. 5,621,859), and further in view of Gupta et al. (US Patent No. 6,138,095).

Russell et al. fail to teach a system comprising table look-up means for setting the transition bias in accordance with the number of phonetic segments in the utterance, and direct setting means for setting the transition bias as proportional or equal to the number of phonetic segments in the utterance.

However, a system comprising "table look-up means for setting the transition bias" is well known in the art as evidenced by Schwartz et al., which disclose a lookup-table where transition probabilities are stored for each transition from each grammar state to each possible following word (see column 15, lines 15-18 and 27-29; see also Figure 8). Motivation for the combination would be to reduce the amount of computation done by the system by storing transition probabilities already calculated.

Both Russell and Schwartz et al. fail to teach setting the transition bias in accordance with, or proportional to, the number of phonetic segments in the utterance. However, setting the transition bias in accordance with, or proportional to, the number of phonetic segments in the utterance would have been obvious to one of ordinary skill in the art given the invention by

Gupta et al.. Gupta et al. disclose that rejecting performance of speech recognition can be improved if a different rejection threshold is selected for each utterance length (see column 3, lines 46-48), which is a synonymous to the idea of setting different transition biases that is utterance-length dependent or proportionally dependent, which includes setting the bias equal to the length. Gupta et al. teach that this would improve recognition performance for different utterance lengths (see column 1, line 58, through column 2, line 3).

Russell, Schwartz and Gupta et al. do not explicitly teach setting the transition bias in accordance with, or proportional to, the number of phonetic segments in the word. However, this feature is well known in the art as evidenced by Gandhi et al. in col. 7, ll. 25-33, which teaches segmenting test utterances into words and computing a duration [number of phonetic segments] normalized likelihood score for each word in the input string.

It would have been obvious for one of ordinary skill in the art at the time the invention was made to bias the transition probabilities in dependence on the estimated of duration of a word instead of an utterance so that recognition performance for relatively long and/or short words can be improved.

Conclusion

13. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure:

Bahl et al. teaches "Speech Recognition with Hidden Markov Models of Speech Waveforms."

Kushner et al. (US Patent 5,617,509) teaches a method, apparatus, and radio optimizing Hidden Markov Model speech recognition.

14. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Eunice Ng whose telephone number is 571-272-2854. The examiner can normally be reached on Monday through Friday, 8:30 a.m. - 5:00 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on 571-272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

EN
02/15/07

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER